Research report 170

LINEAR FEATURE EXTRACTION USING THE MULTIRESOLUTION FOURIER TRANSFORM

ANDREW DAVIES, ROLAND WILSON
(RR170)

A hierarchical space-time image representation - the Multiresolution Fourier Transform (MFT) - is discussed, and its properties are used in order to define a spectrum based model for the class of simple linear features, such as straight lines and edges. The model is used to derrive a method of identifying theses features and estimating their parameters, i.e. position and orientation. Results are presented for this process using a test image.

The effect of white noise is considered, and a correlation measure is defined to show the effect of the noise upon the model. A method of further improving the results for a noisy image, by smoothing with oriented ellipses is also considered.

Department of Computer Science
University of Warwick
Coventry CV4 7AL
United Kingdom

November 1990

Contents

List of Figures

Introduction

Image analysis is concerned with using computers to analyse and process images. Starting from a digitised image, processing operations must be performed that transform the initial array of grey levels, into some "higher level" representation. When choosing the transformation it is necessary to decide what exactly we wish to do with the image. Do we simply wish to enhance it or remove noise so that it is of "better quality", by some objective or subjective criteria? Do we wish to segment the image, that is break it up into regions based on some measure of similarity between pixels such as grey level or other statistical properties? Do we wish to detect features in the image such as edges or straight lines, or even textures? Often we need a combination of these.

The purpose of this work is to look at extracting straight lines and edge features from a grey level image, using the Multiresolution Fourier Transform representation [5] of the image.

Section 1 looks at feature extraction from both a non-hierarchical and a multiresolution point of view. The need for multiresolution is discussed, and the reason for a need to compromise between how accurately you can say both what a feature is, and where it is (the uncertainty principle[23]) is considered. Section 1.2.2 looks at one attempt to tackle this problem - the Wavelet transform.

The next section discusses the Multiresolution Fourier Transform (or MFT), which is used as a basis for the following work. This is a representation which allows information to be combined from levels which have high frequency resolution, but low spatial resolution, and levels which have high spatial resolution, but low frequency resolution.

Section 3 describes the model used for the feature extraction process. This model is based on the properties of the spectrum of a single linear feature (edge or line), eg. energy orientation and linear phase relationship. Results are then presented for this method of feature extraction at different levels of the MFT.

The effects of noise are considered in section 4. The feature extraction algorithm is run on a noisy image with no other filtering. Next, ways of improving these results by oriented smoothing of the MFT coefficients are examined. An oriented smoothing scheme is defined and the results of using it on a noisy image are presented.

1 Feature Extraction

One major area of image analysis is concerned with extracting "features" of some sort from a given image. For example a program may be used to detect lines and edges and to produce an edgemap of the image. In all cases a set of desired features must be decided upon before processing of the image starts. The different approaches to this

problem may be broadly classified into two sets. First there are the non-hierarchical approaches, which operate over the whole of the image, either in the spatial domain or some transform domain, and produce a representation of the features contained therein. The other approach is to break down the image in some way and then examine the individual parts of it. Such an approach allows different areas of the image to be considered at different levels of detail or at different resolutions.

1.1 Non-hierarchical Methods

1.1.1 The Hough Transform

The Hough Transform [10] transforms the image from the spatial domain into a domain parameterised by the constants used to define the feature of interest. Although the transform may be generalised to detect more complex features (e.g. quadratics, circles or ellipses), only its use with straight lines is considered here.

The input to a Hough transform is usually an edge map of the image under consideration. A simple way to obtain such an edge map is to use a spatial mask, which can be convolved with the image [1, 12]. After the convolution the image points of highest intensity will be those positioned on a feature of interest. Using different kernels, lines at different orientations may be detected. Other kernels may be used for detecting edges and, as described in [12] a set of orthonormal masks can be used to detect all edges and lines within an image.

As pointed out in [10] it is preferable to consider the normal parameterisation of straight lines, as this avoids the problems of infinite gradient associated with a cartesian parameterisation. Suppose a straight line in the image (x, y) plane is defined by

$$\rho = x \cos\theta + y \sin\theta \tag{1}$$

where ρ is the distance of the line from the origin, and θ is the orientation of the line line with respect to the x-axis. This line may be represented by the parameters (ρ, θ). This representation is the Hough Transform of the line.

Consider an image $I(x, y)$. Any point $I(x_i, y_i)$ may have, in the continuous case, an infinite number of different straight lines running through it, each at a different orientation. When considering a discrete version of the transform space, there are only M distinct orientations that the line may have, and it may be shown ([10]) that each single point in the image space transforms to a sinusoidal curve in the transform space.

The Hough Transform thus produces a histogram in the parameter space. The maxima of this histogram correspond to those lines which are most appropriate to the points in the edge map.

The Hough transform may be adapted to a hierarchical approach ([19]). In this work the image is split into a number of small sub-images, for which the Hough

transform is used to find small line segments. These segements are then propagated up through the hierarchy and grouped together to obtain larger segments.

1.1.2 Space - Frequency Representations

Of the many space-frequency representations available the most common are the Gabor representation [11], the Short Time Fourier Transform [18] and the Wigner distribution [7, 8, 9]. A review of these representations may be found in [6]. Each of these approaches, however, considers the image at a single scale. The scale in the STFT representation is determined by the window function used. The properties of this window function in space and frequency must be decided upon before processing starts. Because of this the scale must be consistent over the entire image.

1.2 Hierarchical Methods

1.2.1 Advantages of Multiresolution Representations

A multiresolution approach to image analysis allows different regions of the image to be analysed in the way that best suits their feature structure. Within an image it is highly unlikely that all the features of interest will be of the same size, so it seems natural to consider these different features at different scales.

Scale-space approaches, as discussed in [14, 24], allow the image to be viewed at differing levels of spatial-resolution, and hence at different levels of detail.

Amongst the most common scale-space methods used in image processing are those that provide a trade off between frequency domain resolution and spatial domain resolution. This is necessary because of the uncertainty principle, [21],[17], which prevents an image being considered at arbitrarily high resolutions in both domains simultaneously.

A multiresolution approach, in comparison to the space-frequency methods described above, allows the image to be considered at different scales and the information at these different scales combined or compared. No decision has to made about the scale at which features in the image can be best examined needs to made beforehand - the decisions can be made during processing and the scale can be different for different parts of the image.

1.2.2 The Wavelet Transform

The Wavelet Transform ([13],[15],[16]) of a 1-dimensional signal is defined, as the convolution of the signal with a "wavelet" kernel by

$$Wf(s,u) = \int_{-\infty}^{+\infty} f(x)\sqrt{s}\psi(s(x-u))dx \qquad (2)$$

and is effectively a decomposition of $f(x)$ onto a set of functions $\psi_s(x-u)$, where

$$\psi_s(x) = \sqrt{s}\psi(sx) \tag{3}$$

and s is a scale factor. This set of functions is called a "wavelet family" and $\psi_s(x)$ is a "wavelet". The wavelet family consists entirely of translations and dilations of the single function $\psi_s(x)$. The transform may also be viewed as bandpass filtering.

The advantage of the wavelet transform over the STFT is that the scale parameter s determines the resolution in the spatial and frequency domains. The wavelets have the property that if we define the standard deviation of energy (about 0) for $\psi(x)$ in the spatial domain as σ_u, then the energy of $\psi_s(x-u_0)$ is concentrated about u_0 with standard deviation σ_u/s. Similarly if we define $\hat{\psi}(\omega)$,as the Fourier transform of $\psi(x)$, then the energy of $\hat{\psi}_s(\omega-\omega_0)$ is concentrated about $s\omega_0$ with standard deviation $s\sigma_\omega$. Thus it can be seen that increasing s increases the concentration in the spatial domain, while spreading out the wavelet in the frequency domain. Conversely in order to increase concentration in the frequency domain, s must be reduced. Therefore (see [16]) the wavelet has a "resolution cell" in the phase-space equal to

$$[u_0 - \frac{\sigma_u}{s}, u_0 + \frac{\sigma_u}{s}] \times [s\omega_0 - s\sigma_\omega, s\omega_0 + s\sigma_\omega]. \tag{4}$$

The next stage is to consider the 2-dimensional version of the wavelet transform. This may be defined for signal $f(x,x) \in L^2(R^2)$, as

$$Wf(s,(u,v)) = \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)s\Psi(s(x-u),s(y-v))dxdy. \tag{5}$$

Again this may be considered as filtering with a bandpass filter of impulse response $\Psi(s(x-u),s(y-v))$. The value of s determines the amount of bandpass smoothing that is performed on the image.

Mallat's work with the wavelet transform [15, 16] has been based on using wavelets as an orthonormal basis for computing the level of detail present at each level of a multiresolution representation of an image. Following Witkin, Mallat defines the detail at a level as that information which is present at that level, but not at any level of coarser resolution. In doing this he provides a description of the original, sampled image, which consists of a coarse representation of the image, at some low level of resolution, and an orthogonal set of levels of detail that must be added back to the image, as its resolution is increased, in order to recover the original image. Since the detail levels are orthogonal the representation is non-redundant (cf. Laplacian pyramid [4]), and only N^2 samples are needed to represent an $N \times N$ image.

The restrictions of the Wavelet representation become apparent when you consider the resolution cell of the 1-dimensional case. If you wish to consider some feature at high frequency then the frequency resolution is low. For example, if we wish to look at frequencies centered on ω_0 then the resolution cell is

$$[u_0 - \sigma, u_0 + \sigma_u] \times [\omega_0 - \sigma_\omega, \omega_0 + \sigma_\omega]. \tag{6}$$

If we then wish to look at the frequencies around $2\omega_0$, s, the scale parameter, must be made equal to 2. This gives the following resolution cell

$$[u_0 - \frac{\sigma_u}{2}, u_0 + \frac{\sigma_u}{2}] \times [2\omega_0 - 2\sigma_\omega, 2\omega_0 + 2\sigma_\omega]. \quad (7)$$

It can be seen that although the frequency information is now centered on $2\omega_0$, the frequency resolution has been halved and the spatial resolution doubled.

In the case of the two-dimensional transform, if the Wavelet functions are considered in terms of angle and radial frequency, then the angular resolution is fixed in the frequency domain. A "long" line (or edge) segment has a narrow oriented spectrum, whereas a shorter segment has a spectrum that is less concentrated. It is surely desirable to be able to alter angular resolution to take account of this.

2 The Multiresolution Fourier Transform

The MFT, or Multiresolution Fourier Transform (see [5]), provides a grid of local spectral estimates, one for each of a number of spatial regions of the image. A single level of it may be seen as a variant of the STFT, using finite prolate spheroidal sequences (FPSS's) as windowing functions. FPSS's are bandlimited functions, with maximal concentration of energy in the spatial domain and are described in [20, 22].

There are a number of different levels of the MFT, with the bottom level being the spatial representation of the image and the top one being the DFT of the image. The levels in between give a combination of both spatial and frequency information, with the spatial resolution decreasing from lower to higher levels, and the frequency resolution increasing (see figure 1).

In this work the relaxed MFT [5] is used. This is a redundant representation, as for each level of the MFT there are four times as many coefficients as there are in the original image. Level n of the MFT of an $N \times N$ image, is a grid of $N/2^{l-1} \times N/2^{l-1}$ spectra, each of size $2^n \times 2^n$. If we consider a 128×128 image then level 4 will be a 16×16 array of spectra, each of which will have 16×16 coefficients. At level 5 there will be 8×8 spectra, each of 64×64 coefficients (see figure 2).

Since there are $N/2^{l-1} \times N/2^{l-1}$ spectra, then each one may be considered as representing the Fourier transform of the corresponding $2^n \times 2^n$ spatial block. The spatial block is assumed to be have its origin between the 4 central coefficients of the block.

In the calculation of the MFT used for this work, each MFT block is centred so that the origin is between the centre 4 coefficients, and therefore the coefficients are numbered as shown in figure 3. In order to achieve this centering, it is necessary to centre the DFT of the image before calculation of the MFT is started. This centering, or shifting of the DFT by 0.5 of a sample, in the frequency domain causes a π phase jump at the edge of the image in the spatial domain. The phase jump is effectively a

strong edge feature, the strength of which depends on the grey levels at the edge of the image. This edge effect is caused as the shift in the DFT domain is introduced by multiplying the image by $e^{-j\pi(x+y)/N}$ in the spatial domain. Since the top edge is at $y = -N/2$ and the bottom at $y = N/2$, the top pixels at the top edge are multiplied by $e^{\frac{-j\pi x}{N}}e^{\frac{-j\pi}{2}}$ while those on the bottom edge are multiplied by $e^{\frac{-j\pi x}{N}}e^{\frac{+j\pi}{2}}$. Since the pixel values are initially real, this introduces a π phase difference.

The coefficients of level l of the MFT of an $N \times N$ image are defined by the equation

$$Mf(l,u,v,x,y) = \sum_{x'}\sum_{y'} f(S_l x - x', S_l y - y')g_l(x', y', S_l, \Omega_l)e^{j2\pi\Omega_l(x'u+y'v)} \tag{8}$$

where

$$0 \leq x, y < \Omega_l \tag{9}$$

$$u = -\frac{S_l - 1}{2}, -\frac{S_l - 3}{2}, ..., -0.5, 0.5, ..., \frac{S_l - 3}{2}, \frac{S_l - 1}{2} \tag{10}$$

$$v = -\frac{S_l - 1}{2}, -\frac{S_l - 3}{2}, ..., -0.5, 0.5, ..., \frac{S_l - 3}{2}, \frac{S_l - 1}{2} \tag{11}$$

and Ω_l is the bandwidth of the FPSS, S_l is the size of the region into which the spatial energy of the FPSS is concentrated. The 2-d relaxed FPSS, $g_l(x, y, S, \Omega_l)$ is cartesian separable and defined by

$$g_l(x, y, S, \Omega_l) = g_l(x, S_l, \Omega_l)g_l(y, S_l, \Omega_l) \tag{12}$$

At a level l

$$S_l = 2^l \tag{13}$$

and

$$\Omega_l = \frac{N}{2^{l-1}} \tag{14}$$

Since this work is concerned with the extraction of lines and edges, which have relatively broad bandwidths, it was decided to that some method was required to reduce the effects of low frequencies, that would otherwise dominate the MFT spectra. In the work by Calway [5] this was done by "pre-whitening" the image, i.e. emphasising the higher frequencies, while reducing the strength of the lower frequencies, of the image before the feature extraction process. In this work the MFT is generated using a high pass FPSS filter in order to remove the low frequency emphasis. The FPSS used is equivalent to that used to produce the MFT and hence the stop band of the filter is dependent upon the level of the MFT and so it effectively cuts out the 'dc' from each block, and reduces the magnitude of the centre 4 coefficients.

The MFT gives a trade-off between space and frequency resolution. At the higher levels the frequency resolution is higher, while at the low level there is a higher spatial resolution. The level number, l, plays the part of the scale parameter in the wavelet representation. At any given level the frequency resolution is constant for all

frequencies, there is not the logarithmic resolution that is associated with the wavelet transform.

It is possible infact to consider the MFT as a generalisation of the Wavelet transform, since it is possible to obtain many, non-equivalent Wavelet transforms by combining coefficients from the MFT. Alternatively it can be seen as a generalisation of the STFT to account for uncertainty in scale.

3 Linear Feature Extraction

If we assume that each MFT block is a good approximation to the spectrum of the relevant spatial block, then we can use the Fourier properties of the MFT block to determine the contents of the spatial block. Indeed the "non-relaxed" MFT can be seen as an ideal estimate of the spectrum [5]. In this work the features of interest are assumed to be linear (i.e. straight lines or edges) and run either completely, or part way, through the block. Thus by searching a MFT block for those properties it would have if a linear feature were present in the spatial block it is possible to determine whether a significant feature exists in the spatial block, and if so its position and orientation.

3.1 Feature Model

It can be shown [5] that if an image has a linear feature running all, or part, of the way through it, and nothing else of significant energy, then the energy in its spectrum will be orientated orthogonally to the feature orientation, as shown below in figure 4, where ρ represents the distance of the centre of mass of the feature from the origin, and θ is the orientation of the feature with respect to the x-axis. An image feature

$$f(x, y) = f(x\cos\theta + y\sin\theta) \tag{15}$$

transforms to

$$F(u, v) = 2\pi F(u\cos\theta - v\sin\theta)\delta(u\cos\theta + v\sin\theta) \tag{16}$$

The calculation of this orientation is described in section 3.1.1.

However this orientation of the spectrum is not sufficient to indicate a single linear feature. Multiple parallel features or oriented textures will also have a spectrum which is concentrated along an orthogonal orientation. It is therefore necessary to consider another property of an oriented linear feature, i.e. the linear relationship between the phases of the spectrum coefficients. This is discussed in section 3.1.2 along with its use in determining the position of the feature.

3.1.1 Orientation

If it assumed that a block contains a single linear feature, the orientation of the feature can be obtained by calculating the orientation of maximum concentration of the frequency domain energy. Although this may be done by considering the energy at a number of discrete orientations, as in [5], in this work the energy from each point in a modified MFT block was used to produce a moment of inertia tensor [3]. The reason for, and nature of, this modification is described below, and the energy value used defined in equation (42).

The modified block, which has the same energy orientation as the original MFT blocks, may be considered as having energy in an ellipse, centered on the origin, where the energy at a given point (u, v) within the block is $E(u, v)$. The inertia matrix $\mathbf{I}$, is defined by

$$\mathbf{I} = \begin{bmatrix} I_{00} & I_{01} \\ I_{10} & I_{11} \end{bmatrix} \tag{17}$$

where

$$I_{00} = \sum_u \sum_v \frac{E(u,v)u^2}{r_{uv}} \tag{18}$$

$$I_{11} = \sum_u \sum_v \frac{E(u,v)v^2}{r_{uv}} \tag{19}$$

$$I_{01} = I_{10} = -\sum_u \sum_v \frac{E(u,v)uv}{r_{uv}} \tag{20}$$

and

$$r_{uv} = \sqrt{u^2 + v^2} \tag{21}$$

The r_{uv} factor is included as it was decided to reduce the greater emphasis to energy further away from the origin.

The aim of this calculation is to find the orientation of the major and minor axes of the ellipse. If the calculation of the tensor had been made in a coordinate system defined by these axes, then I_{01} and I_{10} would equal 0. Therefore the major and minor axes of the ellipse are the principle axes of the tensor([3]), and are represented by its eigenvectors. If the orientation of maximum energy concentration, $\theta_{\max}$, is defined as the orientation of the major axis of the ellipse, then this is indicated by the direction of the eigenvector, λ_1, corresponding to the largest eigenvalue of the tensor $\mathbf{I}$, $\mathbf{e_1}$. If

$$\mathbf{e_1} = \begin{bmatrix} e_{10} \\ e_{11} \end{bmatrix} \tag{22}$$

then $\theta_{\max}$ is given by

$$\theta_{\max} = \arctan\left(\frac{e_{11}}{e_{10}}\right) \tag{23}$$

Since it can be easily shown that

$$\frac{e_{11}}{e_{10}} = \frac{\lambda_1 - I_{00}}{I_{01}} \tag{24}$$

then the orientation can be calculated from

$$\theta_{\max} = \arctan\left(\frac{\lambda_1 - I_{00}}{I_{01}}\right) \tag{25}$$

where

$$\lambda_1 = \frac{I_{00} + I_{11} + \sqrt{I_{00} + {I_{11}}^2 - 4(I_{00}I_{11} - {I_{01}}^2)}}{2} \tag{26}$$

The elements of the inertia tensor can be related to the second order partial derivatives of the image (see Bigün [2]). Thus the tensor, and hence the orientation, could be calculated in the spatial domain using the gradient of the image at each point.

3.1.2 Position

The other attribute of a linear feature is its position, i.e. the distance of its centre of mass from the origin. In [5] this is considered as the single distance from the origin, ρ in figure 4, whereas in this work it is considered as the two separate x and y displacements. If there is a simple linear feature in an image then its MFT coefficients will have a linear phase relationship, dependent upon the position of the feature in the image.

Papoulis([17]) shows that the center of gravity η of a one dimensional, real continuous signal is related to the derivative of its Fourier transform, by the equation

$$\eta = -\phi'(\omega) = -\frac{d}{d\omega}\phi(\omega)\bigg|_{\omega=0} \tag{27}$$

where $\phi(\omega)$ is the phase of the Fourier transform. If it is assumed that for signal of interest the phase can be modeled by a linear equation, i.e.

$$\phi(\omega) = -(\eta\omega + c) \tag{28}$$

then the value of the derivative will be the constant η, and therefore the center of gravity will be a distance $-\eta$ from the origin.

Extending this to the two dimensional case, then due to the separability of the Fourier transform the displacements of the center of gravity in the x and y directions can be related to the partial derivatives of the transform phase,

$$\eta_x = -\frac{\partial}{\partial u}\phi(u,v)\bigg|_{u=0,v=0} \tag{29}$$

$$\eta_y = -\left.\frac{\partial}{\partial v}\phi(u,v)\right|_{u=0,v=0} \tag{30}$$

If it is again assumed that for linear features of interest the phase can be modeled by a linear combination of u and v, i.e.

$$\phi(u,v) = -(\eta_x u + \eta_y v + c) \tag{31}$$

then the partial derivatives of the phase will each be constant and equal to the displacements of the center of gravity.

Considering an MFT block as a discrete estimate of a spectrum which can be modelled as above

$$\frac{\phi_M(u+1,v) - \phi_M(u,v)}{\Delta u} \approx \left.\frac{\partial}{\partial u}\phi(u,v)\right|_{u=0,v=0} = -\eta_x \tag{32}$$

$$\frac{\phi_M(u,v+1) - \phi_M(u,v)}{\Delta v} \approx \left.\frac{\partial}{\partial v}\phi(u,v)\right|_{u=0,v=0} = -\eta_y \tag{33}$$

where $\phi_M(u,v)$ is the phase of coefficient (u,v) in the MFT block.

The phase relationships in u and v directions may therefore be modelled by a linear equation. If the magnitudes of the MFT are modelled as exponentials, then the coefficients can be modelled by the linear recursion equations

$$s(u,v) = \alpha_u s(u-1,v) + w_u(u,v) \tag{34}$$

and

$$s(u,v) = \alpha_v s(u,v-1) + w_v(u,v) \tag{35}$$

where in each case $|\alpha_u|$ and $|\alpha_v|$ (both of which are assumed to be in the interval $[0,1]$) determine the shape of the two exponentials and

$$\arg(\alpha_u) = -\eta_x \tag{36}$$

$$\arg(\alpha_v) = -\eta_v \tag{37}$$

The innovation term allows for imperfections in the model and may be considered as white noise, i.e

$$E\{w(u_1,v_1)w^*(u_2,v_2)\} = 0, \text{if } u_1 \neq u_2 \text{ and } v_1 \neq v_2 \tag{38}$$

and

$$E\{w(u,v)s^*(u-1,v)\} = 0 \tag{39}$$

It should be noted that the values of η represent the distance from the origin within a block that is $N \times N$, and are in the interval $[-\pi,\pi]$. The actual physical displacement of the center of gravity is given by the expression, $(N/2\pi)\eta$.

The above equations may be considered as a first order Markov model of the spectrum of a concentrated linear feature, as each coeficent is dependent only on its immediate neighbour and a white noise innovation term.

Considering a block of MFT coefficients as described in section 2, the transform energy is spread out, due to the windowing used in the calculation of the MFT, so that there are three, or more, rows of coefficients parallel to the orientation of maximum energy concentration. Within this band of coefficients, it may be assumed that the the Markov model described above holds. Thus the phase relationship in the x-direction can be found by the following method.

Defining a block of MFT coefficients (see equation (8)),

$$B_{xyl}(u,v) = Mf(l,u,v,x,y) \tag{40}$$

a new block of coefficients, $\mathbf{\Delta X}$, can be obtained from $\mathbf{B}$, based on the following equations:

$$\Delta X(u,v) = \begin{cases} 0.5(B^*(u-1,v)B(u,v) + B^*(u,v)B(u+1,v)) & \text{if } 0 < u < N-1 \\ 0 & \text{otherwise} \end{cases} \tag{41}$$

Both $\mathbf{\Delta X}$ and $\mathbf{B}$ may be considered as N-square arrays. A similar process may be followed for the phase differences in the y direction to produce a block of coefficients $\mathbf{\Delta Y}$. It should be noted that the energy used to form the inertia tensor derived in section 3.1.1 is defined by the equation:

$$E(u,v) = 0.5(|\Delta X(u,v)|^2 + |\Delta Y(u,v)|^2) \tag{42}$$

Once this block has been computed the phase difference, ϕ_x, can be found by averaging over the block, weighting according to the magnitudes of the coefficients. Since the energy is concentrated along the orientation $\theta_{\max}$, it was decided to concentrate only on a narrow band of coefficients along that orientation. Thus the following weighting function was used:

$$b(u,v) = \begin{cases} 1 & \text{if } |d(u,v)| \leq 2 \\ 0 & \text{otherwise} \end{cases} \tag{43}$$

where $d(x,y)$ is the distance of the coefficient $B(x,y)$ from a straight line running through the origin, at orientation $\theta_{\max}$, and is defined by

$$d(u,v) = u\cos(\theta_{\max}) - v\sin(\theta_{\max}) \tag{44}$$

If we define R_x by

$$R_x = \frac{1}{N^2}\sum_{u=0}^{N-1}\sum_{v=0}^{N-1}\Delta X(u,v)b(u,v) \tag{45}$$

The value of R_x is effectively the average of the coefficients within a narrow band along the orientation of the energy, and can therefore be analysed in terms of the Markov model defined above.

$$\begin{aligned} E\{\Delta X(u,v)\} &= 0.5E\{(B^*(u-1,v)B(u,v) + B^*(u,v)B(u+1,v))\} \\ &= 0.5E\{(B^*(u-1,v)B(u,v)\} + 0.5E\{B^*(u,v)B(u+1,v))\} \\ &= 0.5E\{(s^*(u-1,v) + w^*(u-1,v))(s(u,v) + w(u,v))\} \\ &\quad +0.5E\{(s^*(u,v) + w^*(u,v))(s(u+1,v) + w(u+1,v))\} \\ &= 0.5[E\{s^*(u-1,v)s(u,v)\} + E\{s^*(u,v)s(u+1,v)\}] \\ &= 0.5\alpha_u[E\{s^*(u-1,v)s(u-1,v)\} + E\{s^*(u,v)s(u,v)\}] \end{aligned}$$

If the variance of the feature is denoted σ_s^2, this simplifies to

$$E\{\Delta X(u,v)\} = \alpha_u \sigma_s^2 \tag{46}$$

This gives, since σ_s^2 is real,

$$\arg(R_x) = \arg(\alpha_u) = \eta_x \tag{47}$$

Due to the spatial and frequency sampling,

$$-\frac{N}{2} \leq \delta_x, \delta_y \leq \frac{N}{2}. \tag{48}$$

The spatial blocks however are separated by only $N/2$ due to the number of MFT blocks. Therefore any feature positioned between two block centres will be "seen" by both blocks, one from either side. In order that no feature is seen twice, if a feature is found to be more than $N/4$ away from the spatial origin of the block,i.e. if

$$|\arg(R_x)| > \frac{\pi}{2} \text{ or } |\arg(R_y)| > \frac{\pi}{2} \tag{49}$$

then it is assumed to be in some other block and ignored.

It can be seen that $|R_x|^2$ and $|R_y|^2$ are dependent upon the of the amount of energy in the orientation of the feature, and on how well the feature fits the linear phase model. Given the same amount of energy in the feature orientation, $|R_x|$ and $|R_y|$ will be greatest when the phases diferences between adjacent coefficients are constant over the area of energy concentration. By combining them a value can be obtained that gives a measure of the magnitude of the particular feature:

$$E = 0.5(|R_x|^2 + |R_y|^2) \tag{50}$$

3.1.3 Analysis of Signal Moments

Assuming that the signal is "clean", and that it fits the Markov model, the first and second order moments after the phase differencing can be analysed.In the MFT block the signal is assumed to have the following statistical properties:

$$E\{s(u,v)\} = 0 \tag{51}$$

$$E\{s(u,v)s^*(u,v)\} = \sigma_s^2 \tag{52}$$

If $\mathbf{\Delta S}$ denotes the block of differenced coefficients, i.e.

$$\Delta S(u,v) = 0.5[s(u,v)s^*(u-1,v) + s^*(u,v)s(u+1,v)] \tag{53}$$

then its first and second order moments can be calculated as follows:

$$\begin{aligned} E\{\Delta S(u,v)\} &= 0.5E\{s(u,v)s^*(u-1,v) + s^*(u,v)s(u+1,v)\} \\ &= 0.5E\{\alpha_u s(u-1,v)s^*(u-1,v) + s^*(u,v)\alpha_u s(u,v)\} \\ &= \alpha_u \sigma_s^2 \end{aligned}$$

Defining

$$R_{\Delta S} = E\{\Delta S(u,v)\Delta S^*(u,v)\} \tag{54}$$

$$\begin{aligned} R_{\Delta S} &= 0.25E\{[s(u,v)s^*(u-1) + s^*(u,v)s(u+1,v)] \\ &\quad [s^*(u,v)s(u-1) + s(u,v)s^*(u+1,v)]\} \\ &= 0.25[E\{s(u,v)s^*(u-1)s^*(u,v)s(u-1,v)\} \\ &\quad + E\{s(u,v)s^*(u-1,)s(u,v)s^*(u+1,v)\} \\ &\quad + E\{s^*(u,v)s(u+1)s^*(u,v)s(u-1)\} \\ &\quad + E\{s^*(u,v)s(u+1,v)s(u,v)s^*(u+1,v)\}] \end{aligned}$$

By expanding fourth order terms into sums of second order terms (see [17, chapter 12]),

$$\begin{aligned} R_{\Delta S} &= 0.25[2\sigma_s^4 + 6\sigma_s^4|\alpha_u|^2 + \sigma_s^4|\alpha_u|^2(E^2\{s^2(u,v)\} + E^2\{s^{*2}(u,v)\}) \\ &\quad + \alpha_u^2 E\{s^{*2}(u,v)\}E\{s^2(u-1,v)\} \\ &\quad + \alpha_u^{*2} E\{s^2(u,v)\}E\{s^{*2}(u-1,v)\}] \end{aligned}$$

Assuming that the real and imaginary parts of the signal coefficients have the same statistics and noting the hermitian symmetry of the transform:

$$E\{s^2(u,v)\} = 0 \tag{55}$$

The expression simplifies to:

$$R_{\Delta S} = 0.5\sigma_s^4[1 + 3|\alpha|^2] \tag{56}$$

This gives the mean-square of ΔS, which must satisfy

$$R_{\Delta S} \leq 2\sigma_s^4 \tag{57}$$

due to $|\alpha|^2$ being in the interval $[0, 1]$.

As the magnitudes of the MFT coefficients are being modelled by exponentials, and the phase by a linear relationship, its value will depend upon the resolution of the MFT blocks. At one level in the MFT, the frequency resolution in each direction is twice that at the previous level. This suggests that if the value of α at level l is denoted α_l, then the value at level $(l+1)$ should be

$$\alpha_{l+1} = \sqrt{\alpha_l} = \sqrt{|\alpha_l|} e^{\frac{\arg(\alpha_l)}{2}} \tag{58}$$

Since $|\alpha|$ is less than 1, this implies that the effective value of $R_{\Delta S}$ increases with level.

3.2 Results for Clean Images

The processes described in sections 3.1.1 and 3.1.2 were tried on different levels of the MFT for the girl image, figure 5.

This feature extraction process was used for each block, in a given level of the MFT, to find the position, orientation and strength of the features within it. Each block was assumed to have one linear feature. Once this had been done then an image showing the features was generated. Figures 6 - 9 show the results for drawings made from using levels 3 - 6 of the girl image.

In order to facilitate the display of these images on a laser printer, the magnitude of the features have been thresholded, with those of magnitude less than $\sim 4\%$ of the maximum being ignored. The rest are drawn as black lines.

It may be seen that in some of these results there is an feature parallel to the edge of the image. This is due to the MFT being centred between 4 coefficients as described in section 2. In order to reduce the effects of these "image-edge features", any features in the edge blocks that are parallel to the appropriate image edge are assumed to be due to this effect and therefore ignored. If these edges are strong, then there will be ghosts of these edges that appear further into the image.

One cause of this problem may be the lowpass filter used in the generation of the MFT. As noted earlier this filter is the same that is used to produce the MFT windowing. The size of the filter causes features to be spread to a size comparable with that of the MFT blocks. Thus the edges, which are very strong features, contribute a noticeable amount of energy to the neighbouring blocks inside the image. One way around this is to increase the size of the lowpass frequency window and therefore reduce the spreading of features in the spatial domain. This was done for level four of the image and the feature extraction results are shown in figure 10. In this instance the lowpass filter was an FPSS with a bandwidth double that originally used.

4 Noise

4.1 Analysis of Noise Moments

This section contains an analysis of the coefficients of the differences blocks $\mathbf{\Delta X}$ and $\mathbf{\Delta Y}$, when noise is present in the image. Assume that the noise in the MFT coefficients has the following properties:

$$E\{n(u,v)\} = 0 \tag{59}$$

$$E\{n(u,v)n^*(u,v)\} = \sigma_n^2 \tag{60}$$

$$E\{n(u,v)n^*(u \pm 1,v)\} = E\{n(u,v)n^*(u,v \pm 1)\} = E\{n(u,v)n^*(u \pm 1,v \pm 1)\} = \beta\sigma_n^2 \tag{61}$$

The correlation between the adjacent noise samples is caused by the overlapping of adjacent frequency windows in the relaxed MFT. By considering these windows as raised cosines an approximation of β can be found of about 0.254. If $\mathbf{\Delta N}$ is used to denote the differenced noise coefficients the following moments can be expressed

$$\begin{aligned} E\{\Delta N(u,v)\} &= 0.5E\{n(u,v)n^*(u-1) + n^*(u,v)n(u+1,v)\} \\ &= 0.5E\{n(u,v)n^*(u-1)\} + 0.5E\{n^*(u,v)n(u+1,v)\} \\ &= \beta\sigma_n^2 \end{aligned}$$

If we define

$$R_{\Delta N} = E\{\Delta N(u,v)\Delta N^*(u,v)\} \tag{62}$$

then

$$\begin{aligned} R_{\Delta N} &= 0.25E\{[n(u,v)n^*(u-1) + n^*(u,v)n(u+1,v)] \\ &\quad [n^*(u,v)n(u-1) + n(u,v)n^*(u+1,v)]\} \\ &= 0.25[E\{n(u,v)n^*(u-1)n^*(u,v)n(u-1,v)\} \\ &\quad + E\{n(u,v)n^*(u-1,)n(u,v)n^*(u+1,v)\} \\ &\quad + E\{n^*(u,v)n(u+1)n^*(u,v)n(u-1)\} \\ &\quad + E\{n^*(u,v)n(u+1,v)n(u,v)n^*(u+1,v)\}] \end{aligned}$$

By expanding the fourth order terms and substituting for known values

$$\begin{aligned} R_{\Delta N} &= 0.25[2\sigma_n^2 + 6\beta^2\sigma_n^2 + 2|E\{n(u,v)n^*(u-1,v)\}|^2 \\ &\quad + E\{n^{*2}(u,v)\}E\{n(u-1,v)n(u+1,v)\} \\ &\quad + E\{n^2(u,v)\}E\{n^*(u-1,v)n^*(u+1,v)\}] \end{aligned}$$

Assuming that real and imaginary terms in the noise are uncorrelated this simplifies to

$$R_{\Delta N} = 0.5\sigma_n^4[1 + 3\beta^2] \tag{63}$$

The variance of $\Delta \mathbf{N}$ can therefore be expressed as

$$\sigma^2_{\Delta N} = 0.5\sigma_n^4[1+\beta^2] \tag{64}$$

The above analysis considers the block differencing along the u direction of the MFT block, and hence relates to the calculation of the x direction spatial displacement of the image feature. Similar analysis can be used for the v direction differencing.

Although this considers noise in isolation it cannot be used in the consideration of a noisy signal as it does not take into account all the correlation effects introduced by the differencing process. If $x(u,v)$ represents a noisy image, defined by:

$$x(u,v) = s(u,v) + n(u,v) \tag{65}$$

then we can calculate the mean for the MFT block coeficients for the noisy image, and for the differenced coefficients:

$$\begin{aligned} E\{x(u,v)\} &= E\{s(u,v)\} + E\{n(u,v)\} \\ &= 0 \end{aligned}$$

$$\begin{aligned} E\{\Delta X(u,v)\} &= E\{0.5[x(u,v)x^*(u-1,v)\} + E\{x^*(u,v)x(u+1,v)]\} \\ &= 0.5[E\{s(u,v)s^*(u-1,v)\} + E\{n(u,v)n^*(u-1,v)\} \\ &\quad +0.5[E\{s^*(u,v)s(u+1,v)\} + E\{n^*(u,v)n(u+1,v)\}] \\ &= 0.5[\alpha\sigma_s^2 + \beta\sigma_n^2 + \alpha\sigma_s^2 + \beta\sigma_n^2] \\ &= \alpha\sigma_s^2 + \beta\sigma_n^2 \end{aligned} \tag{66}$$

Next an energy measure for the noisy image can be defined as:

$$\begin{aligned} R_{\Delta X} &= E\{\Delta X(u,v)\Delta X^*(u,v)\} \\ &= 0.25E\{[x(u,v)x^*(u-1,v) + x^*(u,v)x(u+1,v)] \\ &\quad [x^*(u,v)x(u-1,v) + x(u,v)x^*(u+1,v)]\} \end{aligned}$$

By substituting for $x(u,v)$ and simplifying, this can be reduced to

$$R_{\Delta X} = 0.5[(3\beta^2+1)\sigma_n^4 + (2+3\beta(\alpha+\alpha^*))\sigma_n^2 + (1+3|\alpha|^2)] \tag{67}$$

The phase of α represents the position of features in the range $\pm\pi$. Since there is no reason why features are more likely in one position than another, it seems valid to assume

$$E\{Re(\alpha)\} = E\{Im(\alpha)\} = 0 \tag{68}$$

This simplifies $R_{\Delta X}$ to

$$R_{\Delta X} = (3\beta^2+1)\sigma_n^4 + 2\sigma_n^2\sigma_s^2 + \sigma_s^4(3|\alpha|^2+1) \tag{69}$$

One way to consider the noise is by analysing the correlation between the clean image and the noisy image. Defining the correlation

$$\begin{aligned}\hat{R}_{\Delta XS} &= 0.25|E\{[x(u,v)x^*(u-1,v)+x^*(u,v)x(u+1,v)] \\ &\quad [s(u,v)s^*(u-1,v)+s^*(u,v)s(u+1,v)]\}|\end{aligned}$$

it can be simplified after substituting

$$x(u,v) = s(u,v) + n(u,v) \tag{70}$$

to

$$\hat{R}_{\Delta XS} = 0.5[2\sigma_s^2\alpha^*(\sigma_s^2\alpha + \beta\sigma_n^2) + \sigma_s^4 + |\alpha|^2] \tag{71}$$

As above it can be assumed that

$$E\{Re(\alpha)\} = E\{Im(\alpha)\} = 0 \tag{72}$$

This leads to the simplification

$$\hat{R}_{\Delta XS} = 0.5[\sigma_s^4(3|\alpha|^2+1)] \tag{73}$$

This can be normalised, so that the maximum value is 1, by dividing by $\sqrt{R_{\Delta X}R_{\Delta S}}$ as defined above, giving

$$R_{\Delta XS} = \frac{\sigma_s^4(3|\alpha|^2+1)}{\sqrt{[(3\beta^2+1)\sigma_n^4 + 2\sigma_n^2\sigma_s^2 + \sigma_s^4(3|\alpha|^2+1)]\sigma_s^4[3|\alpha|^2+1]}} \tag{74}$$

which simplifies to

$$R_{\Delta XS} = \frac{\sigma_s^2(3|\alpha|^2+1)}{\sqrt{[(3\beta^2+1)\sigma_n^4 + 2\sigma_n^2\sigma_s^2 + \sigma_s^4(3|\alpha|^2+1)][3|\alpha|^2+1]}} \tag{75}$$

The feature extraction process described in section 3.1 uses the coefficients along a narrow oriented band. If the feature corresponds to the Markov model, then its energy is concentrated along this band and hence the band contains all (or most of) the feature energy. The noise energy however is evenly spread out across the block. Therefore even if there is the same amount of noise energy as feature energy in the whole block, in the narrow band the signal energy is greater than the noise energy.

Suppose, for example, that the feature energy may be considered totally concentrated in a strip b-pixels wide, horizontally oriented through the block center. The effective signal variance within this strip will be higher than the variance when taken over the whole block. The noise and signal variance, within the band may be defined as

$$\sigma_{sb}^2 = \sigma_s^2\frac{N}{b} \tag{76}$$

and

$$\sigma_{nb}^2 = \sigma_n^2 \tag{77}$$

If it is now assumed that the signal in the original MFT coeficients has variance 1, then at higher levels the effective signal variance, σ_{sb}^2, will increase, causing $R_{\Delta XS}$ to increase.

$$R_{\Delta XS} = \frac{\frac{N}{b}(3|\alpha|^2+1)}{\sqrt{[(3\beta^2+1)\sigma_n^4 + 2\sigma_n^2\frac{N}{b} + \frac{N^2}{b^2}(3|\alpha|^2+1)][3|\alpha|^2+1]}} \tag{78}$$

It was noted in section 3.1.3 that $|\alpha|$ increases with level, thus $|\alpha|^2$ also increases with level. This is a further source of improvement in $R_{\Delta XS}$ at higher levels.

In the case where there is no noise present, i.e. $\sigma_n^2 = 0$,

$$\begin{aligned} R_{\Delta XS} &= \frac{\sigma_s^2[3|\alpha|^2+1]}{\sigma_s^2[3|\alpha|^2+1]} \\ &= 1 \end{aligned} \tag{79}$$

4.2 Noise Measurement

When considering a noisy image, it is necessary to have some measurement of how good the results are in comparison with a clean image. The following measure was used, and it provides a measure of correlation between two vectors, (cf equation 75) and if the data fits the previous model precisely then it should give the same values as for equation (75). Since all the feature extraction work is done on blocks of phase differenced coefficients, $\mathbf{\Delta X}$ and $\mathbf{\Delta Y}$, it is these coefficients that are compared in each case.

$$r = \frac{|\sum(\widehat{\Delta X}_{i,j}(u,v)\Delta X^*_{i,j}(u,v))| + \sum|(\widehat{\Delta Y}_{i,j}(u,v)\Delta Y^*_{i,j}(u,v))|}{\sqrt{\sum|\widehat{\Delta X}_{i,j}(u,v)|^2\sum|\Delta X^*_{i,j}(u,v)|^2} + \sqrt{\sum|\widehat{\Delta Y}_{i,j}(u,v)|^2\sum|\Delta Y^*_{i,j}(u,v)|^2}} \tag{80}$$

where all the summations are over i, j, u, v. It should be noted that $\mathbf{\Delta X}_{i,j}$ and $\mathbf{\Delta Y}_{i,j}$ are the blocks of differenced coefficients for the MFT block (i,j), for the clean image. $\widehat{\mathbf{\Delta X}}_{i,j}$ and $\widehat{\mathbf{\Delta Y}}_{i,j}$ are the estimated values of $\mathbf{\Delta X}_{i,j}$ and $\mathbf{\Delta Y}_{i,j}$ obtained from the noisy image.The value of r will be one if the two pairs of compared blocks are equal, but less than one if they are not, i.e.

$$r\begin{cases} = 1 & \text{if } \Delta X_{i,j}(u,v) = \widehat{\Delta X}_{i,j}(u,v), \Delta Y_{i,j}(u,v) = \widehat{\Delta Y}_{i,j}(u,v), \quad \forall i,j,u,v \\ < 1 & \text{otherwise} \end{cases} \tag{81}$$

4.3 Results for Noisy Images

The following results are for the girl image, with added white gaussian noise, such that the signal-to-noise ratio is 0dB. The SNR is defined by

$$\text{SNR} = 10 \log \left(\frac{\sigma_i^2}{\sigma_n^2} \right) \tag{82}$$

where σ_i^2 is the variance of the clean image, and σ_n^2 is the variance of the noise. The girl, with the noise, is shown in figure 11. The first column of table 1 contains the results of equation 80 for various MFT levels. There has been no smoothing or filtering except that which is introduced by producing the phase difference blocks, as described in section 3.1.2, and the high-pass filtering used in the generation of the MFT.

Due to its higher concentration the highpass filtering removes a higher proportion of the signal energy than of the noise energy, which serves to reduce the SNR. The amount of this reduction is however dependant upon level, due to the size of the highpass filter. As the level increases so does the SNR.

In figure 34 the values of r for the noisy image are plotted along with the values of R calculated taking into account the variation of SNR with level. It is assumed that $b = 4$ at all levels but that the SNR decreases with level and that α varies as previously described.Level 3 corresponds to $N = 8$, level 4 to $N = 16$ etc. Both the theoretical and experimental result show that the correlation improves with level. They are not the same since the image does not consist entirely of linear features and hence it does not fit exactly to the feature model used.

It should be noted that the differencing operations have removed a lot of the noise at the higher levels. This implies that using these differenced blocks will provide a robust method of feature extraction.

In section 3.2 the problem of "edge of image" features was partialy addressed by doubling the bandwidth of the lowpass filter used to remove low frequency energy in the image. When this approach is used on a noisy image, however, removing more lowpass energy means removing more of the signal energy than of the noise energy, and hence reduces the SNR. On level 4 the value of the correlation function (defined in equation (80)) becomes 0.509, compared to 0.801 which is obtained using the smaller filter.

4.4 Elliptical Smoothing

The next stage was to consider some form of elliptical smoothing to try and reduce the effects of noise on the feature extraction. Considering the linear phase relationship for a straight line feature, the phase difference coefficients along the orientation of maximum energy concentration, should be constant, (cf. Markov model in section

MFT LEVEL	Values of (σ_x^2, σ_y^2)					
	No Filter	(1.0,0.25)	(2.0,0.5)	(4.0, 0.25)	(4.0,0.5)	(4.0,1.0)
3	0.425	0.481	0.577	-	0.622	-
4	0.801	0.831	0.854	0.846	0.876	0.874
5	0.985	0.988	0.990	0.990	0.991	0.991
6	0.997	0.997	0.998	0.998	0.998	0.998

Table 1: Table of r values for MFT levels of Girl (0dB SNR)

3.1.2). This suggests that smoothing the coefficients along the direction of this orientation, will reduce noise components in the coefficients. To do this each coefficient is replaced by the average of the surrounding coefficients, weighted by an oriented function, in this case an ellipse.

In order to do this smoothing before any processing on the differenced coefficients is performed, it is not known which coefficients lie on the correct orientation, therefore all coefficients are treated equally and averaged along the orientation of a vector from the origin to the coefficient. The weighting function for the averaging is an ellipse defined by

$$\omega(u,v) = \exp\left(-\frac{(u')^2}{2\sigma_u^2} - \frac{(v')^2}{2\sigma_v^2}\right) \tag{83}$$

where

$$\begin{bmatrix} u' \\ v' \end{bmatrix} = \begin{bmatrix} \cos(\theta_{u,v}) & \sin(\theta_{u,v}) \\ -\sin(\theta_{u,v}) & \cos(\theta_{u,v}) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \tag{84}$$

$$\sin(\theta_{u,v}) = \frac{u}{\sqrt{u^2+v^2}} \tag{85}$$

$$\cos(\theta_{u,v}) = \frac{-v}{\sqrt{u^2+v^2}} \tag{86}$$

The values of σ_u^2 and σ_v^2 are used to define the shape of the elliptical weighting function. If $\sigma_u^2 = \sigma_v^2$ then the function will be circular symmetric, if not then the values will determine the ratio of the length of the major axis to the length of the minor axis.

4.5 Results after Elliptical Smoothing

The elliptical smoothing was performed for the 0dB girl image, figure 11 for different levels of the MFT, and with different values of σ_x^2 and σ_y^2. The values of the r correlation function for these experiments is shown in table 1.

Since the smoothing process causes the energy in the MFT blocks to be spread out, the comparisons used to produce these results, come from comparing smoothed

versions of the noisy image with smoothed versions of the clean image. This means that the central energy in both cases is similarly spread out.

The image features derived from the feature extraction phase, both with and without elliptical filtering are shown in figures 12 - 33.

The elliptical smoothing process, when the right size of ellipse is chosen, does improve the correlation. If the signal component of the differenced coefficients was constant and the noise components white, averaging would remove the noise. However its effectiveness is restricted by the actual signal it operates on. The noise becomes correlated by both the MFT process and the phase differencing, and the signal is not constant as has been shown above. The gain caused by the elliptical smoothing is therefore reduced.

5 Conclusions/Further Work

This work has considered the use of the Multiresolution Fourier Transform, or MFT, for extracting linear features from images. A model of these features has been defined and an algorithm developed which uses this model to detect appropriate features. This algorithm has been run on both clean and noisy images and it results presented. In the case of the noisy image even with no further filtering the results, especially at the higher levels, show the method to be robust in the presence of noise.

The next step was to consider some way of reducing the effects of noise, especially at lower levels. An oriented smoothing approach was adopted, using an elliptical filter. This was used to perform weighted averaging on each pixel in each block.

The work so far has considered each MFT block as a separate entity, although some measure of the energy concentration along the axis of maximum energy concentration can be used to decide whether there is a single feature present in the block. A further stage of development would be to combine information from neighbouring blocks and different levels. The robust feature estimates at higher level, for example, could be used to help the lower level feature estimates where noise is more of a problem.

Another development would be to generalise the feature model to include simple curves, such as quadratics or circular arcs. At most levels the eyes of the girl image are not detected due to their high curvature, and in order to detect them some model of circular features within an MFT block would be needed.

References

[1] D. H. Ballard and C. M. Brown. *Computer Vision.* Prentice Hall, New Jersey, 1982.

[2] J. Bigün. Local symmetry Features in Image Processing. Linköping Studies in Science and Technology Disertation No. 179, Computer Vision Laboratory, Linköping University, Sweden, 1988.

[3] A. I. Borisenko and I. E. Tarapov. *Vector and Tensor Analysis.* Dover Publications, Inc., New York, 1979.

[4] P. J. Burt and E. H. Adelson. The Laplacian pyramid as a compact image code. *IEEE Trans. Comp.*, COM-31:532–540, 1983.

[5] A. Calway. *The Multiresolution Fourier Transform: A general Purpose Tool for Image Analysis.* PhD thesis, Department of Computer Science, The University of Warwick, UK, September 1989.

[6] A. Calway and R. G. Wilson. Hierarchical Descriptors for Nonstationary 1 and 2 Dimensional Signal Processing. Research Report RR108, Department of Computer Science, University of Warwick, UK, October 1987.

[7] T. A. C. M. Classen and W. F. G. Mecklenbräuker. The Wigner Distribution - A Tool for Time-Frequency Analysis. Part I: Continuous-Time Signals. *Phillips J. Research*, 35:217–250, 1980.

[8] T. A. C. M. Classen and W. F. G. Mecklenbräuker. The Wigner Distribution - A Tool for Time-Frequency Analysis. Part II: Discrete-Time Signals. *Phillips J. Research*, 35:276–300, 1980.

[9] T. A. C. M. Classen and W. F. G. Mecklenbräuker. The Wigner Distribution - A Tool for Time-Frequency Analysis. Part III: Relations with other Time-Frequency Signal Transformations. *Phillips J. Research*, 35:372–389, 1980.

[10] R. O. Duda and P. E. Hart. Use of the Hough Transform To Detect Lines and Curves in Pictures. *Communications of the ACM*, 15(1):11–15, January 1972.

[11] D. Gabor. Theory of Communications. *Proc. IEE*, 93(26):429–441, November 1946.

[12] R. C. Gonzalez and P. Wintz. *Digital Image Processing.* Addison Wesley Publishing Company, Inc., 1987.

[13] A. Grossman and J. Mortlet. Decomposition of Hardy functions into square integrable wavelengths of constant shape. *SIAM J. Math. Anal.*, 15:723–736, 1984.

[14] T. Lindeberg. Scale-Space for Discrete Signals. *IEEE Trans. P.A.M.I*, 12(3):234–254, March 1990.

[15] S. G. Mallat. A Theory for Multiresolution Signal Decomposition: The Wavelet Representation. *IEEE Trans. P.A.M.I*, 11(7):674–693, July 1989.

[16] S. G. Mallat. Multifrequency Channel Decompositions of Images and Wavelet Models. *IEEE Trans. A.S.S.P*, 37(12):2091–2110, December 1989.

[17] A. Papoulis. *Signal Analysis.* McGraw-Hill, 1984.

[18] M. R. Portnoff. Time-frequency Representation of Digital Signals and Systems Based on Short-Time Fourier Analysis. *IEEE Trans. A.S.S.P*, 28(1):55–69, February 1980.

[19] J. Princen, J. Illingworth, and J. Kitler. A Hierarchical Approach to Line Extraction Based on the Hough Transform. *Computer Vision, Graphics, and Image Processing*, 52(1):57–77, October 1990.

[20] R. G. Wilson. Finite Prolate Spheroidal Sequences and Their Applications I: Generation and Properties. *IEEE Trans. P.A.M.I*, 9(6):787–795, November 1987.

[21] R. G. Wilson and G. H. Granlund. The Uncertainty Principle in Image Processing. *IEEE Trans. P.A.M.I*, 6(6):758–767, November 1984.

[22] R. G. Wilson and M. Spann. Finite Prolate Spheroidal Sequences and Their Applications 2: Image Feature Description and Segmentation. *IEEE Trans. P.A.M.I*, 10(2):193–203, March 1988.

[23] R. G. Wilson and M. Spann. *Image Segmentation and Uncertainty.* Pattern Recognition and Image Processing Series. Research Studies Press Ltd, 1988.

[24] A. P. Witkin. Scale-Space Filtering. In *Proc. of the 8th Int. Joint Conf. on Artificial Intelligence*, pages 1019–1022, 1983.

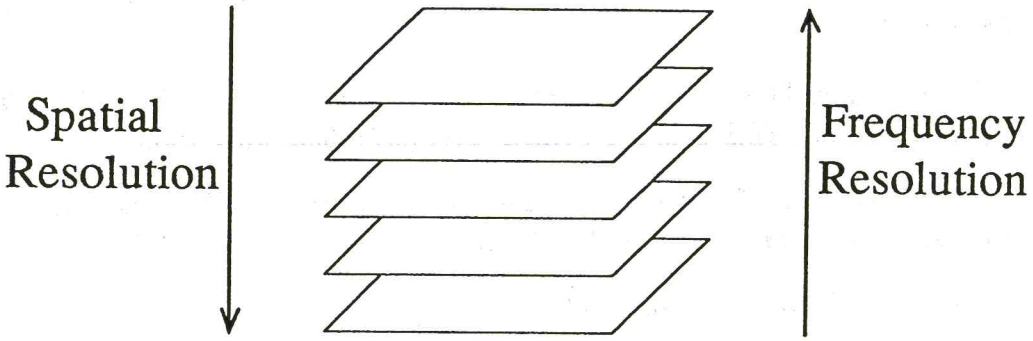

Figure 1: MFT - Levels

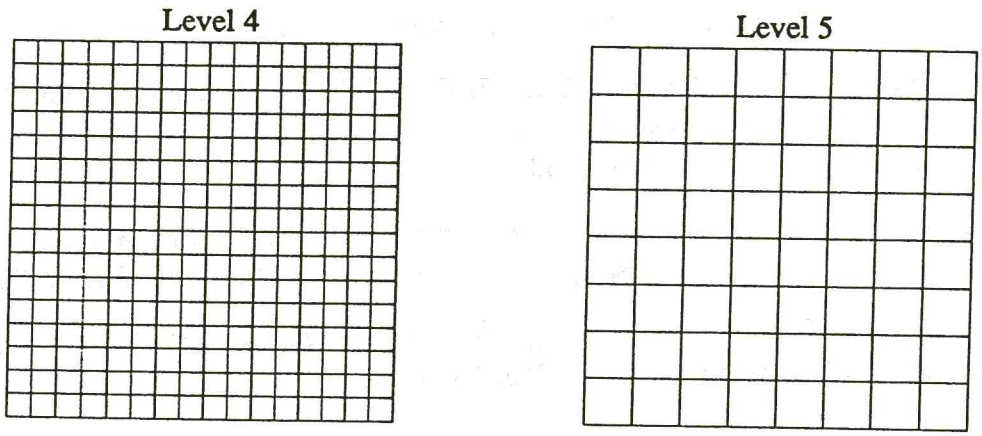

Figure 2: MFT - Levels

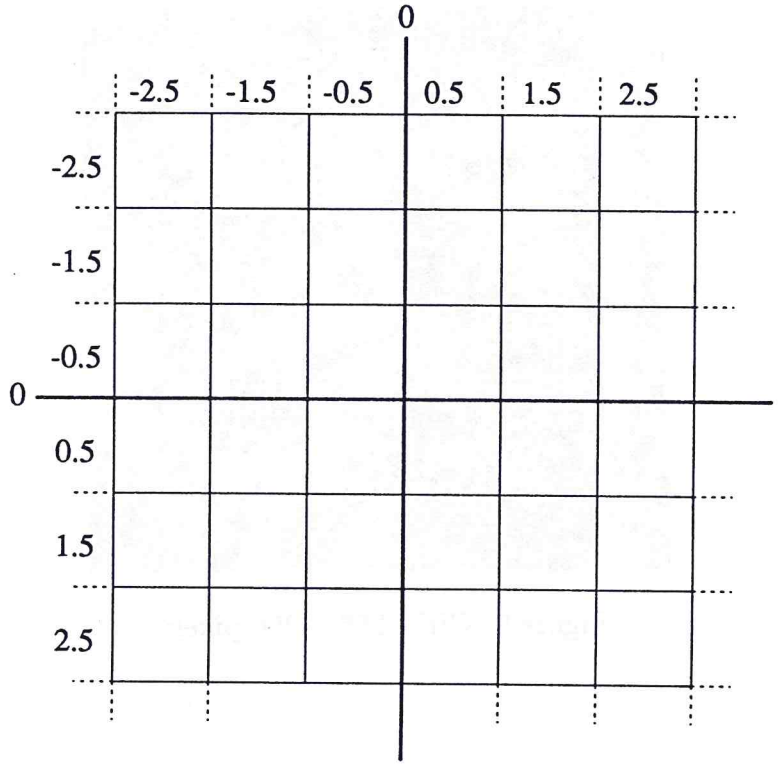

Figure 3: Centre of one MFT block

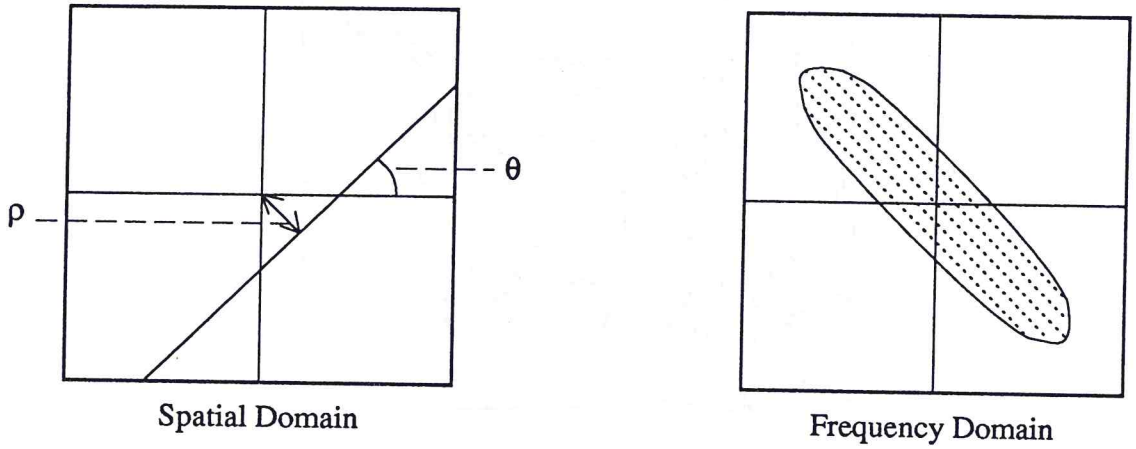

Figure 4: Orientation - Spatial and Frequency Domains

Figure 5: Girl - 256 × 256 pixels

Figure 6: Girl from level 3

Figure 7: Girl from level 4

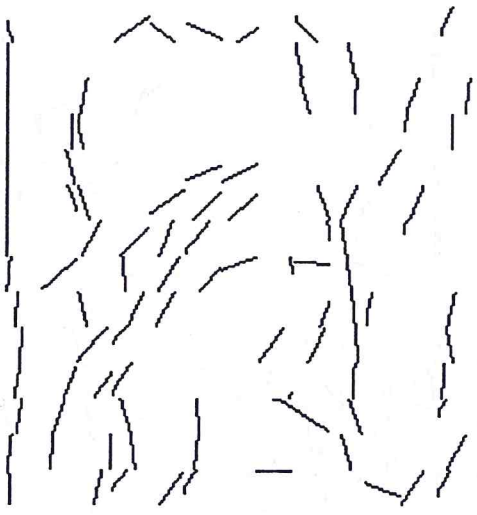

Figure 8: Girl from level 5

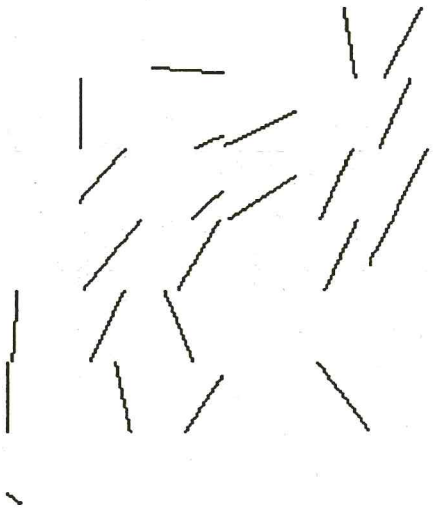

Figure 9: Girl from level 6

Figure 10: Girl from level 4, using larger lowpass filter for MFT

Figure 11: Girl with Noise - 0dB SNR

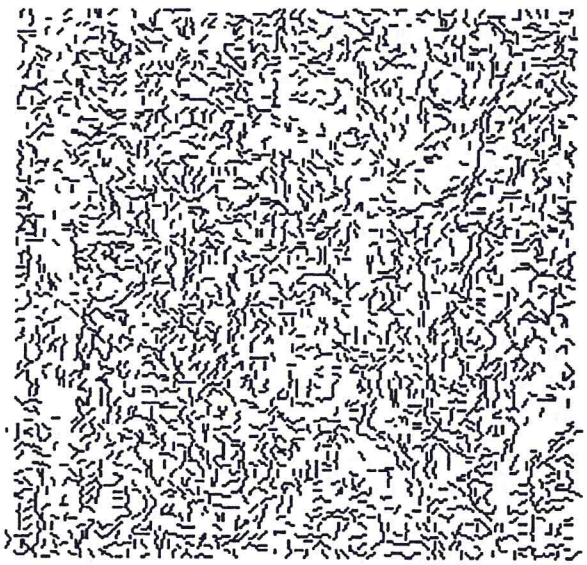

Figure 12: Features from noisy girl, level 3, no elliptical smoothing

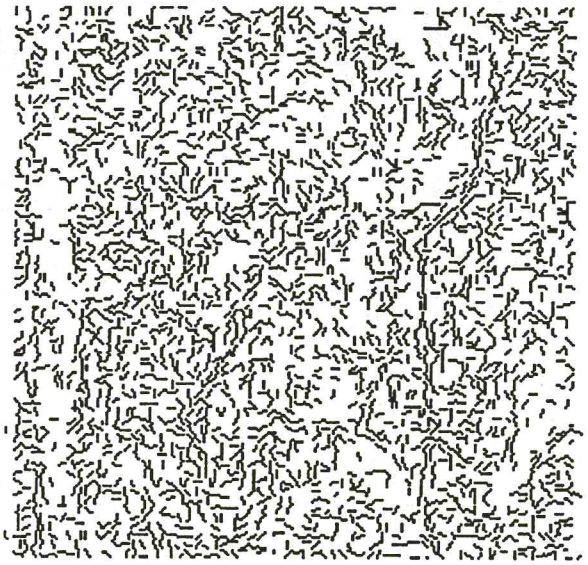

Figure 13: Features from noisy girl, level 3, $(\sigma_x^2, \sigma_y^2) = (1.0, 0.25)$

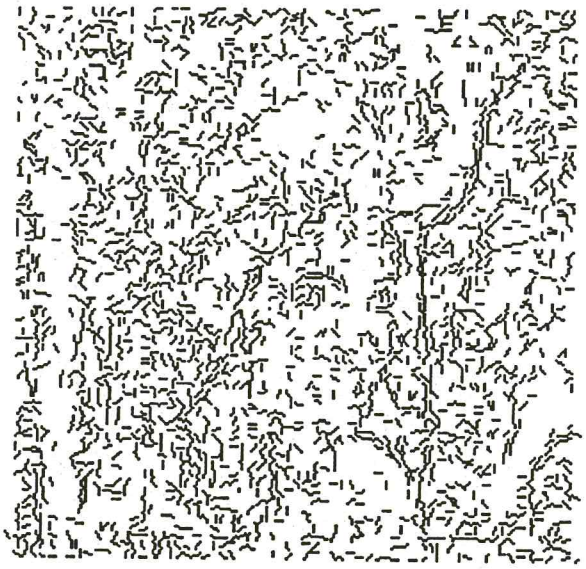

Figure 14: Features from noisy girl, level 3, $(\sigma_x^2, \sigma_y^2) = (2.0, 0.5)$

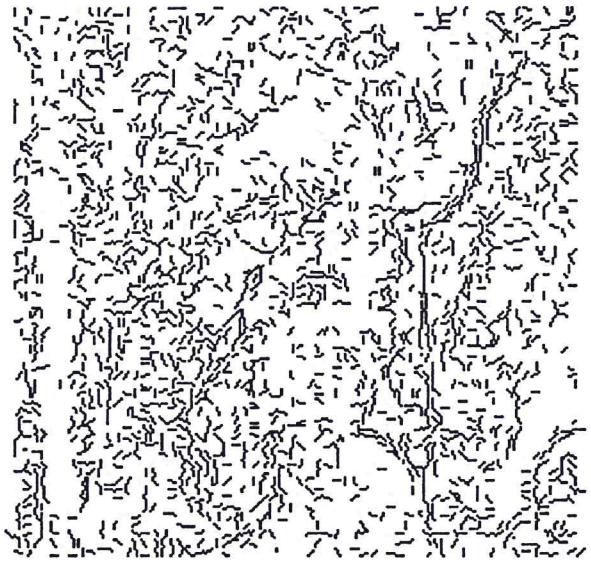

Figure 15: Features from noisy girl, level 3, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.5)$

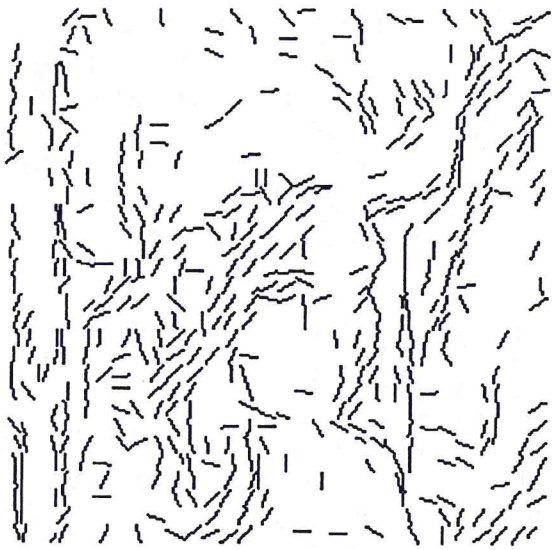

Figure 16: Features from noisy girl, level 4, no elliptical smoothing

Figure 17: Features from noisy girl, level 4, $(\sigma_x^2, \sigma_y^2) = (1.0, 0.25)$

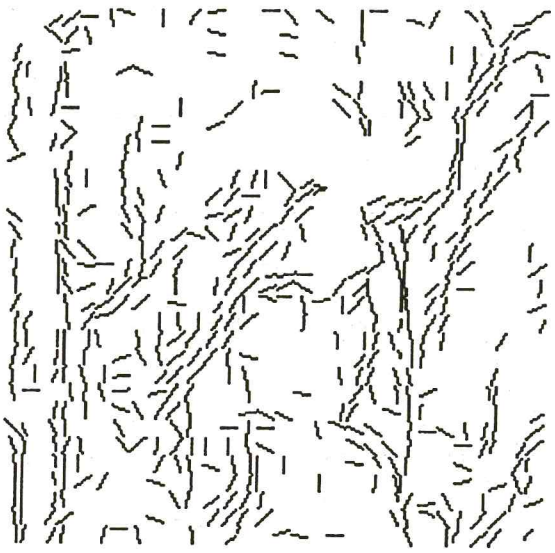

Figure 18: Features from noisy girl, level 4, $(\sigma_x^2, \sigma_y^2) = (2.0, 0.5)$

Figure 19: Features from noisy girl, level 4, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.5)$

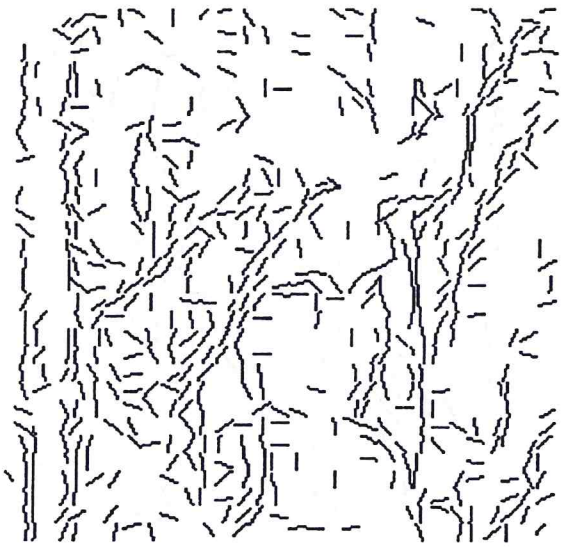

Figure 20: Features from noisy girl, level 4, $(\sigma_x^2, \sigma_y^2) = (4.0, 1.0)$

Figure 21: Features from noisy girl, level 4, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.25)$

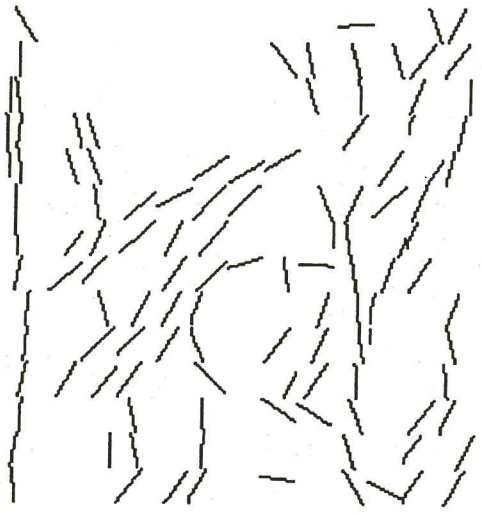

Figure 22: Features from noisey girl, level 5, no elliptical smoothing

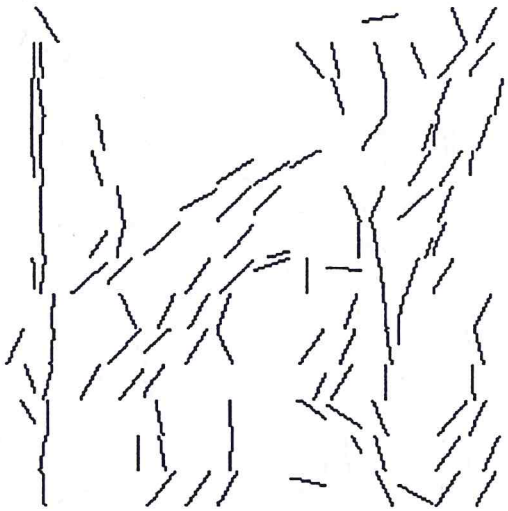

Figure 23: Features from noisy girl, level 5, $(\sigma_x^2, \sigma_y^2) = (1.0, 0.25)$

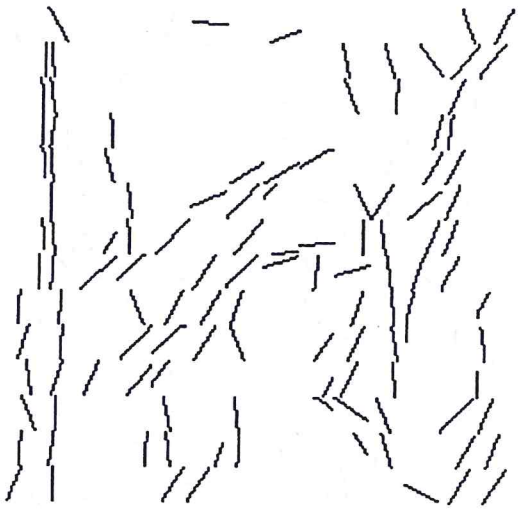

Figure 24: Features from noisy girl, level 5, $(\sigma_x^2, \sigma_y^2) = (2.0, 0.5)$

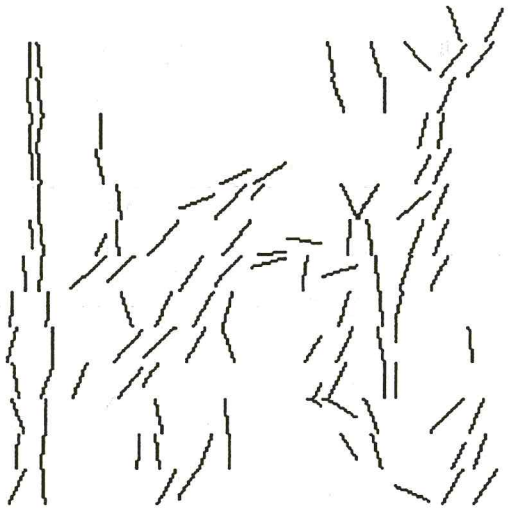

Figure 25: Features from noisy girl, level 5, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.5)$

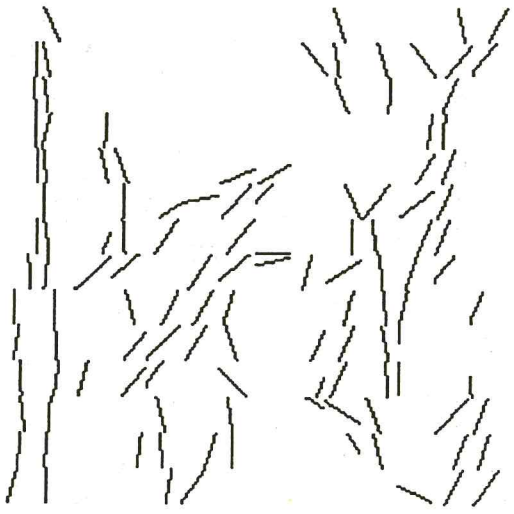

Figure 26: Features from noisy girl, level 5, $(\sigma_x^2, \sigma_y^2) = (4.0, 1.0)$

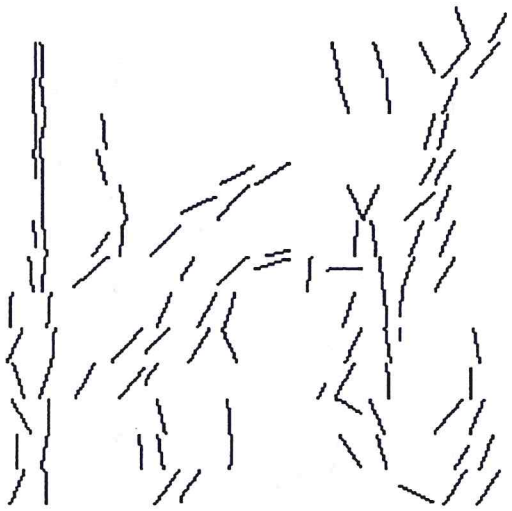

Figure 27: Features from noisy girl, level 5, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.25)$

Figure 28: Features from noisy girl, level 6, no elliptical smoothing

Figure 29: Features from noisy girl, level 6, $(\sigma_x^2, \sigma_y^2) = (1.0, 0.25)$

Figure 30: Features from noisy girl, level 6, $(\sigma_x^2, \sigma_y^2) = (2.0, 0.5)$

Figure 31: Features from noisy girl, level 6, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.5)$

Figure 32: Features from noisy girl, level 6, $(\sigma_x^2, \sigma_y^2) = (4.0, 1.0)$

Figure 33: Features from noisy girl, level 6, $(\sigma_x^2, \sigma_y^2) = (4.0, 0.25)$

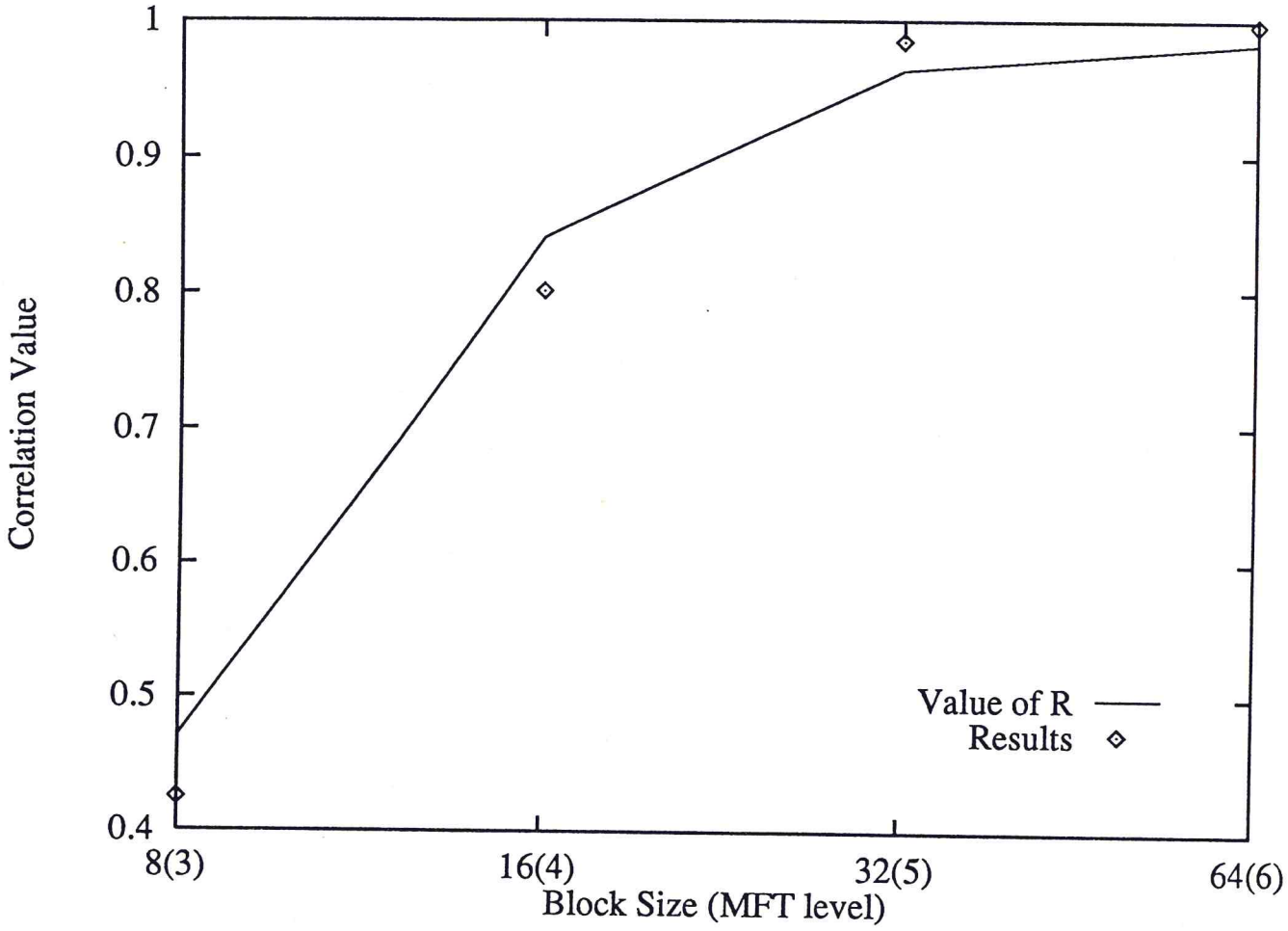

Figure 34: Values of $R_{\Delta XS}$ assuming that the SNR in the orignal MFT decreases with level and that the value of $|\alpha|$ increases with level, and the measured results.